Mechanisms of Covalent Dimerization on a Bulk Insulating Surface

Chunyan Guo[1], Yu Wang[1], Markus Kittelmann[2], Lev Kantorovitch[3*], Angelika Kühnle[2] and Andrea Floris[4*]*

[1]School of Physics and Technology, Wuhan University, Wuhan 430072, China

[2]Institute of Physical Chemistry, University of Mainz, 55099 Mainz, Germany.

[3]Department of Physics, King's College London, London, Strand WC2R 2LS, United Kingdom

[4]School of Mathematics and Physics, University of Lincoln, Brayford Pool, LN6 7TS Lincoln, United Kingdom.

*afloris@lincoln.ac.uk; *lev.kantorovitch@kcl.ac.uk; *yu.wang@whu.edu.cn

CORRESPONDING AUTHORS

Andrea Floris, School of Mathematics and Physics, University of Lincoln, Brayford Pool, LN6 7TS Lincoln, United Kingdom, phone: +44 1522 83 5884.

Lev Kantorovich, Department of Physics, King's College London, London, Strand WC2R 2LS, United Kingdom, phone: +44 020 7848 2160.

Yu Wang, School of Physics and Technology, Wuhan University, Wuhan 430072, China, phone: +86 1587 135 1027.

ABSTRACT

Combining density functional theory and high-resolution NC-AFM experiments, we study the on-surface reaction mechanisms responsible for the covalent dimerization of *4-iodobenzoic acid* (IBA) organic molecules on the calcite (10.4) insulating surface. When annealed at 580K the molecules assemble in one-dimensional chains of covalently bound dimers. The chains have a unique orientation and are the result of a complex set of processes, including a nominally rather costly double dehalogenation reaction followed by dimerization. First, focusing on the latter two processes and using the Nudged Elastic Band method, we analyze a number of possible mechanisms involving one and two molecules and we isolate the key aspects facilitating the reaction on calcite. Second, we show that the insulating surface plays an active role as a catalyst by identifying two relevant processes: one exhibiting an intermediate state of chemisorbed molecules after independent de-halogenations and a second, highly non-trivial exothermic reaction channel where two

iodine atoms "cooperate" to minimize the cost of their individual detachment from the molecules. Both processes have a drastically reduced energy barrier if compared to all other mechanisms analyzed. Knowledge of the formation mechanisms of a covalent assembly on insulators represents an important step towards the realization and control of structures that combine the robustness of covalent architectures with their electronic decoupling from the insulating substrate. This step has potentially important technological applications in nano- and molecular electronics.

1. Introduction

On-surface synthesis is a rapidly growing research field whose goal is the bottom-up fabrication of covalent nanostructures through the coupling of molecular building blocks [1,2,3,4,5], Refs. 6,7 and references therein and Ref. 8. With this strategy, reaction products can be obtained in ultra-high vacuum conditions, without using solvents and thus avoiding problems related to the lack of solubility[6]. Moreover, the surface catalytic properties and the molecular confinement resulting from the presence of a two-dimensional support open a possibility of new reaction pathways towards products not easily obtainable in three dimensions (i.e. in the gas phase or in solutions) [6].

Covalent nanostructures are quite robust in comparison with more "fragile" supramolecular architectures, mediated by weaker interactions (hydrogen bonds, van der Waals and electrostatic [9,10,11,12,13,14,15]), in which the self-assembly is inherently linked to reversible interactions. From the technological point of view, controlling the synthesis of covalent nanostructures on a surface is crucial for their use in molecular devices operating in a wide range of physical conditions. Importantly, the use of *insulating* surfaces guarantees, in addition, the electronic decoupling between the support and the covalent nanostructures. This makes the latter particularly suitable for electronic transport and, in general, for molecular electronic applications like molecular nanodevices[6], as well as for molecular optics [16], electronics and catalysis[17].

These undoubtedly attractive technological perspectives recently stimulated many experiments, combining a large variety of molecular architectures on different substrates[1,2,3,4], Refs. 6-7 and references therein and Refs. 18-25.

Several experimental strategies to synthesize nanostructures on *metal* surfaces have proved very successful. These include the ones exploiting Ullmann-type processes (Refs. 2, 6 and 7 and references therein, Refs. 22, 23), also studied in density functional theory (DFT)-based theoretical studies[7,26]. Or strategies exploiting the activation of specific C-H bonds, with the realization of very selective structures [8, 21, 27, 28], that in some cases have been rationalized by DFT[5]. Or using terminal alkynes (Refs. 6,7 and references therein), or cyclodehydrogenation reactions (Ref. 7 and references therein, Ref. 25), both analyzed theoretically[29,30,31]. These achievements have been reviewed in Refs. 6 and 7 focusing, respectively, on the relevant experimental and theoretical studies to date.

In spite of a considerable progress achieved on metal surfaces, comparatively fewer studies about synthesis on *insulating* substrates are present in the literature, including reactions on bulk surfaces via Ullmann-type processes[3,6,19,32], photochemical initiation[20] and reactions on insulating thin-films (Ref. 6 and references therein, Ref. 18). Several reasons make the on-surface synthesis on insulators rather challenging[3]. First, these substrates cannot be investigated with techniques commonly used for conductive surfaces (e.g., STM, LEED or PES). The second and more fundamental reason is that insulating supports are in general only weakly reactive [3,33,34,35,36] due to a weak hybridization with molecular orbitals (leading to much lower binding energy), to a lower molecule-surface van der Waals interaction, and to their much lower surface energy, if compared to metallic surfaces [3,34]. These facts might result in molecular desorption (before the reaction takes place) and dewetting effects[34], preventing the surface functionalization, as extensively achieved on metals.

An important example of a reactive insulating substrate allowing the assembly of organic molecules is the calcite (10.4) surface [19,32,37,38,39,40,41,42]. This substrate owes its reactivity to an exceptionally high surface energy [3] and to the presence of highly localized charges [34] as active sites suitable for molecular anchoring[3,43].

A number of *covalently* linked structures have been obtained experimentally on calcite (10.4)[6]. Using *4-iodobenzoic acid* (IBA) molecules, molecular dimers have been formed that assemble in one-dimensional chains of side-by-side aligned dimers[4]. More complex molecules have been demonstrated to result in zig-zag and closed ring structures via sequential dehalogenations[19]. These architectures were obtained *via* Ullmann-type or hierarchical sequential linking reactions. However, to the best of our knowledge, there are no theoretical works on the mechanisms of molecular synthesis on insulators providing insights about the hierarchy of possible reaction pathways, although there has been some related work on boron nitride and graphene on Ni (111)[46].

In this work, we rationalize the covalent dimerization mechanisms of IBA molecules within one-dimensional chains observed in non-contact NC-AFM experiments[4] on calcite (10.4) surface when annealed to 580 K. We investigate the role of the surface in the dimerization process, i.e., if it is chemically active or acts only as a two-dimensional support to constrain the molecules on a plane, a necessary condition for them to react. Focusing, in particular, on the conditions making the cleavage of the iodine-phenyl bond feasible, we also address the validity of previous speculations[4, 19] that the cleavage of the iodine-phenyl bond could be facilitated by the deprotonation of the IBA carboxylic group on calcite. Via a detailed analysis based on DFT and the Nudged Elastic Band (NEB) method[44], we propose exothermic mechanisms for of the dehalogenation-dimerization process (whereas the single dehalogenation in the gas phase is strongly endothermic). More importantly, we also identify specific non-trivial mechanisms

where the energy barrier is strongly reduced, rationalizing the experimental observations performed at the relevant temperatures.

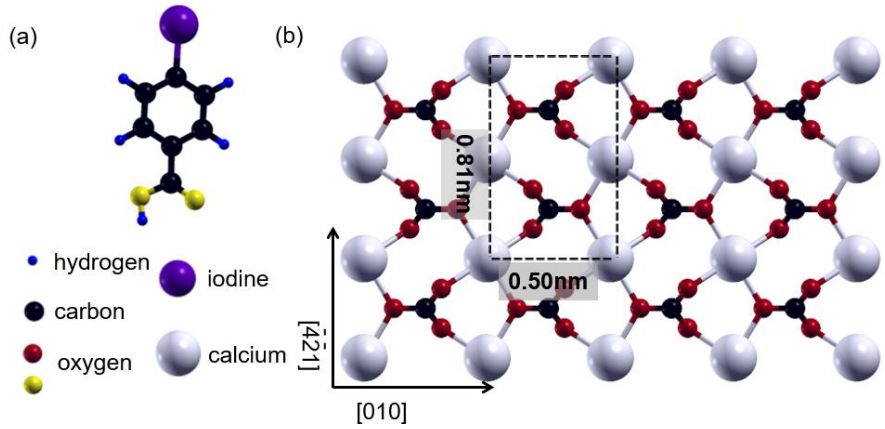


Figure 1. (a) Structure of an IBA molecule in the gas phase. (b) The top layer of the calcite (10.4) surface. The rectangular unit cell (shown by the dashed line) consists of two inequivalent Ca ions and two carbonate groups. The latter are rotated by 180° with respect to each other. For better clarity, oxygen atoms belonging to the molecule and the surface are indicated in yellow and red, respectively.

2. Results and Discussion

Experimental. The IBA molecule and the calcite (10.4) surface are presented in Fig. 1. Our theoretical analysis starts from previous experimental studies based on NC-AFM measurements[4]. The latter shows that upon molecular deposition on the substrate, a sequence of molecular phases is obtained during gradual annealing of the system[4]. When deposited onto a calcite (10.4) surface held at room temperature, molecular islands can be observed at substrate step edges. The islands exhibit a highly ordered inner structure as seen in Fig. 2(a), where we show a zoom into the inner structure. Upon annealing at a moderate temperature of 520 K, a first structural change is observed. From the apparent height of 0.8 nm, we tentatively assigned the molecules to adopt an upright configuration[32], denoted in the following as the *upright* phase. These upright-standing molecules form islands that are shown in Fig 2(b) [32]. Upon annealing to 580 K, a second structural

change is observed. This structural change has previously been assigned to dehalogenation and dimerization of two IBA molecules to form *biphenyl-4,4'-dicarboxylic* acid (BPDCA, for simplicity referred to as "dimer" in the following) [4,39]. The dimers arrange in *chains* oriented along the [-4-21] direction, as shown in Fig. 2(c) [39]. A zoom into one of these chains suggests that it is composed by side-by-side aligned dimers, as shown in Fig. 2(d) (note that the atomic structure of the underlying calcite surface is visible in this image as well). The measured apparent height of the molecules forming the chains (0.3 nm, much smaller than the one in the upright phase, 0.8 nm) *and* the thickness of each chain (compatible with two surface unit cells along [010]), strongly suggest that the chains are formed by flat-lying dimers. Here, the two molecules are connected *via* a covalent bond between two carbon atoms, which are formed upon a dehalogenation reaction between pairs of molecules [4, 39].

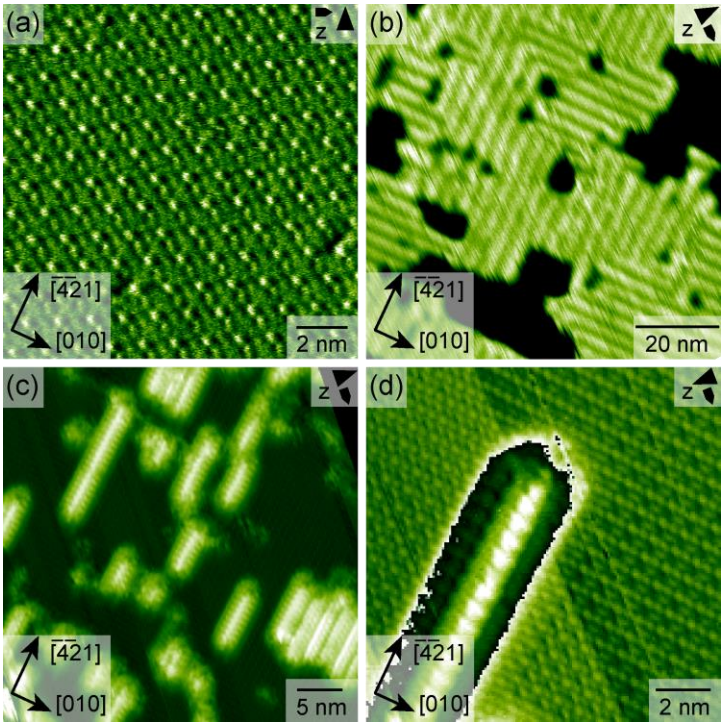


Figure 2. Three arrangements of IBA molecular structures on calcite (10.4) as a function of the sample temperature, obtained with NC-AFM[39]: (a) highly ordered inner structure of molecular islands formed at room temperature; (b) *upright* phase observed after a moderate annealing to 520K; (c, d) chains of dehalogenated IBA dimers arranged along the [$\bar{4}\bar{2}$ 1] direction obtained after annealing to 580 K [4].

IBA molecule in gas-phase. We start our theoretical analysis by simulating a single IBA molecule in the gas phase. The calculated energy cost to dissociate it by removing its iodine atom is 3.19 eV [including two spin channels, see Supporting Information (SI), Sec. S1], compatible with the values to dehalogenate a different but related molecule (iodobenzene) measured experimentally (2.79 eV [45]) and calculated theoretically (3.33 eV) in a recent DFT study [26]. This energy, which in the gas phase also coincides with the corresponding energy barrier, is by far too high for the reaction to happen on the surface, in the relevant interval of annealing temperatures (520-580K) [53], unless a specific reaction mechanism strongly reducing the barrier is at play. Our goal is to investigate the details of this dehalogenation-dimerization reaction, analyzing under which conditions it can happen and hence rationalize the observed covalent assembly on calcite (10.4).

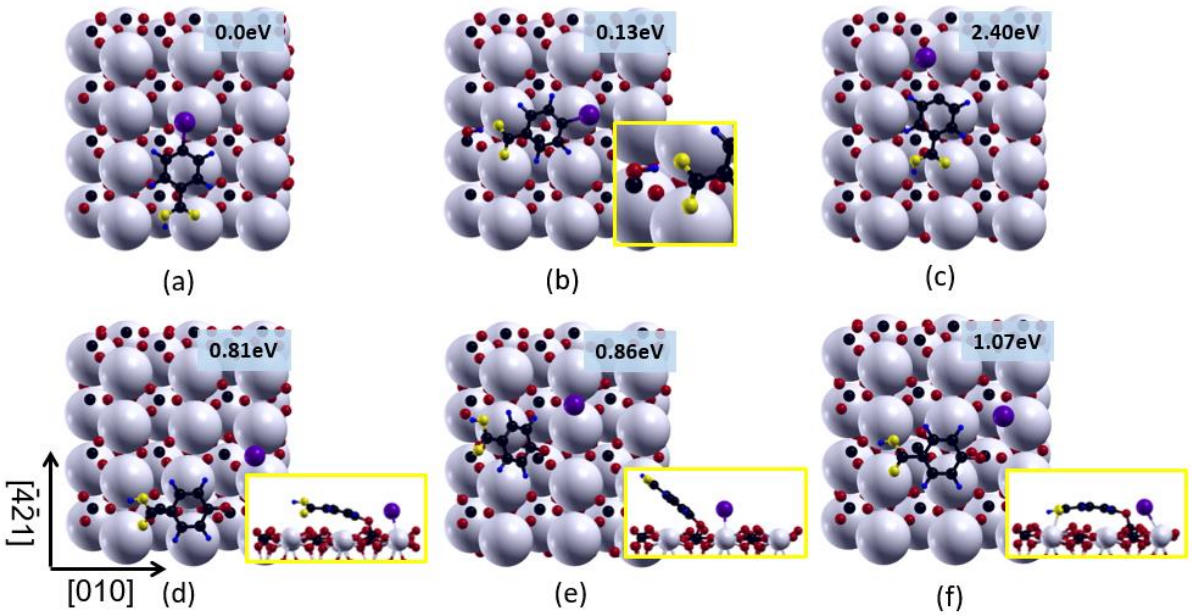


Figure 3. Flat-lying adsorption geometries of a single IBA molecule on calcite (10.4) in its intact (a, b) and dehalogenated (de-Ha) configurations (c-f). (a) The most stable structure found, with the total energy set to 0 eV as a reference for the energies of other configurations. The molecular axis forms an angle of 9^0 with respect to the [$\bar{4}\bar{2}$1] direction. (b) Slightly less stable geometry forming a 19^0 angle with respect to the [010] direction. The cg-H atom relaxes at a position

closer to the s-O than to the IBA carboxylic group oxygen (cg-O) (see inset), so that the role of the s-O and cg-O is interchanged with the cg-H forming an H-bond with the IBA (the calculated cg-H – cg-O distance is ~ 1.4 Å). The two cg-O's bind with the two nearest Ca ions along $[\bar{4}\bar{2}1]$. In order to achieve this configuration, the protruding s-O must be available nearby for the cg-H to bind, with a concomitant breaking of the substrate O-Ca bond. Placing the cg-H far away from the molecule (but still bound to the surface in an equivalent site), implies an energy cost of > 0.75 eV, which corresponds to the H-bond breaking (see Sec. S5). (c) An unstable de-Ha structure with the phenyl radical not directly bonded to a surface atom. (d, e) The most stable de-Ha configurations with the flat phenyl radical bound to a calcite s-O atom. (f) A third, low energy de-Ha structure also forming a bond between the carboxylate group and the surface Ca atom, with a concomitant IBA deformation.

Simulated systems. It can be speculated that the transition between the upright phase and the one in which molecules are organized in chains consists of several distinct elementary processes. An exhaustive *ab initio* analysis of the associated complex energy landscape is rather difficult, as there is no experimental information on the kinetics available that would suggest "reasonable" intermediate states of the system to model. Still, the dehalogenation reaction must be an elementary part of the whole transition, and hence a mechanism leading to a reasonably small energy barrier for this reaction must exist. After making the plausible assumption that the dehalogenation-dimerization is kinetically limited, i.e. it has the largest barrier of all elementary processes and "dominates" the phase transition, we will focus our analysis only on this process. To this end, we will only consider systems made of one single monomer on the surface or a dimer with its axis oriented along [010], or a single chain of equivalent dimers periodically repeated along $[\bar{4}\bar{2}1]$, as shown in Fig. 5 below.

Single monomers on calcite (10.4). Before identifying the factors that might lower the reaction barrier, we show in Figs. 3 and 4 the adsorption geometries of a single monomer on the surface. We considered a large number (~20) of flat-lying (intact and dehalogenated) and upright relaxed configurations (a detailed analysis is presented in Secs. S2 and S3). The most stable structure of a

single molecule is presented in Fig. 3(a), showing a flat molecule on the surface whose adsorption energy is calculated to be 1.53 eV. All energies indicated in Figs. 3 and 4 panels refer to the total energy of this structure. A second, slightly less stable but differently oriented configuration is shown in Fig. 3(b), where in the inset we see that the hydrogen atom of the carboxylic group (denoted cg-H) relaxes to a position closer to the nearest surface oxygen (s-O), away from the IBA carboxylic group oxygen (cg-O). In other words, the cg-H does not detach completely, but still binds to the parent molecule via a hydrogen bond, similarly to the case of a *2,5-dihydroxybenzoic acid* (DHBA) molecule on the same substrate [43].

As for the IBA upright configurations, we see the structure in Fig. 4(a) showing the cg-H still attached to the molecule as in the gas phase, while in the almost degenerate geometries in Figs. 4(b,c) an H-bond similar to that in Fig. 3(b) is realized. The fact that the structures in Figs. 3(a,b) and Figs. 4(a,b,c) are almost degenerate indicates that this slight cg-H detachment can easily happen at finite temperature and that the cg-H atom might actually shuttle from a state to another, due to thermal effects [54].

Relevant de-halogenated (de-Ha) configurations are presented in Figs. 3 (c,d,e,f), with total energies substantially larger than the intact molecules ones. The most stable are shown in Figs. 3(d-f) where the molecule chemisorbs [52] and binds to the surface via its phenyl radical, upon dehalogenation. The structure in Fig. 3(c) (and similar ones, see Sec. S3) where this C-s-O bond is not formed, is considerably more unstable than structures in Fig. 3(d, e) (by > 1.6 eV) and Fig. 3(f) (by >1.3 eV).

Modeling IBA dimers and chains on calcite (10.4). The final state of the reaction analyzed in this work has a dimeric structure, which we investigate by performing several DFT relaxations of conceivable IBA dimers when organized in chains on the surface (Fig. 5). One possible dimerization is realized via a C-C covalent bond upon a double dehalogenation (Fig. 5(a)). The second is realized via H-bonds between the carboxylic groups of two IBA's (like the DHBA

molecule dimers in Ref. 43). We model only dimer configurations along [010], compatible with the experimental observations. The covalent dimer is found to be energetically more stable (by 1.23 eV) than the H-bonded one, which is also substantially larger (by 0.54 nm) than the first, along [010]. Hence, our calculations support the experimental conclusions that the chains are made of covalent dimers[4] and configurations similar to the one in Fig. 5(a) will be considered as the final product of the reaction studied in this work.

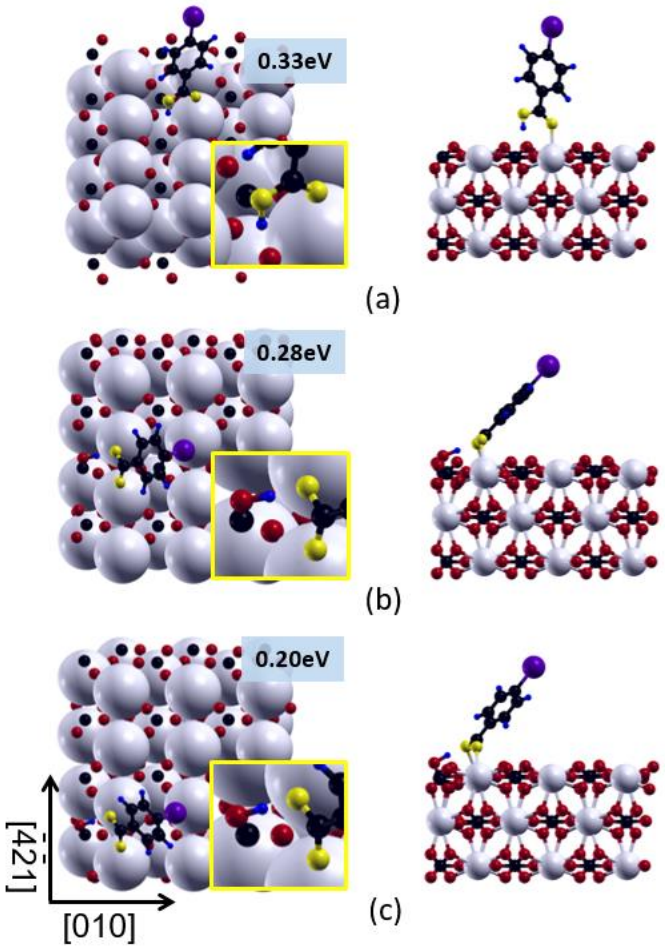


Figure 4. Upright adsorption geometries of a single IBA molecule on calcite (10.4). Left (right) panels give the top (side) views. In (a) the cg-H is attached to the molecule. In (b) and (c), which are a bit more stable, the cg-H is slightly detached from the IBA and forms an H-bond with it (cf. the mechanism mentioned in the caption to Fig. 3 and in Ref. 43). The total energies shown by the structures in the left panels are relative to the energy of the most stable structure of Fig. 3(a).

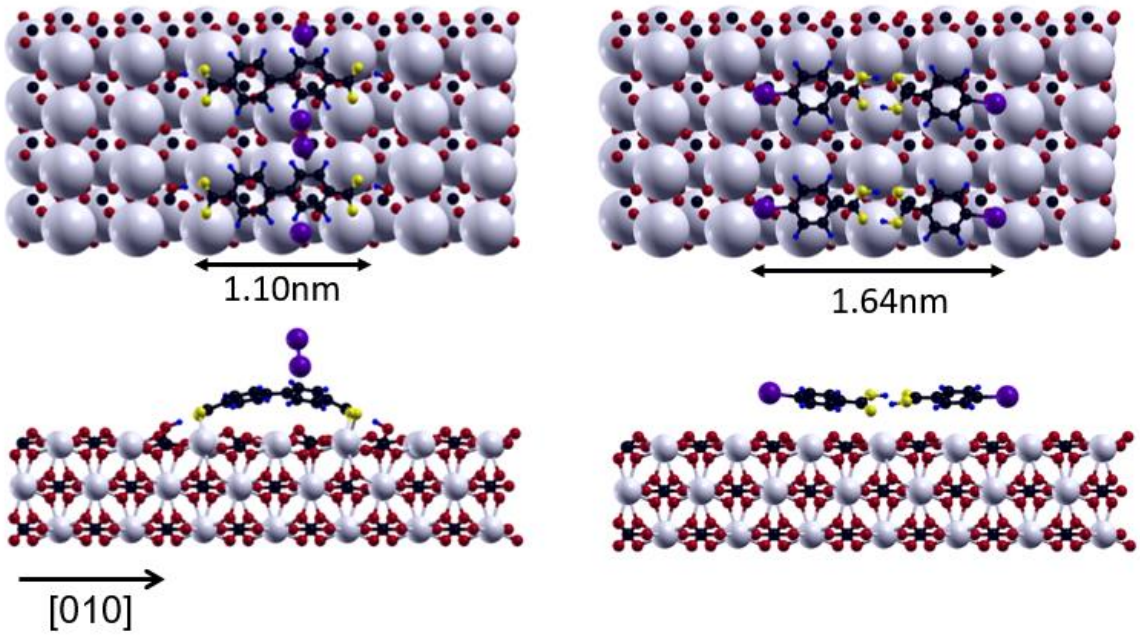


Figure 5. DFT relaxed chains of dimers on calcite (10.4) (top and side views). (a) A chain of covalent dimers made of dehalogenated IBA's connected *via* the C-C links (recombined I_2 molecules are also present in the cell). The side view shows an arc-shaped structure due to a slight mismatch between the length of two calcite unit cells along [010] and the one of the dimer (1.10 nm). The dimer deforms to establish the Ca-O bonds with the surface, a stabilizing factor on the surface. Note that both the H atoms of the carboxylic group (cg-H's) move from the molecules towards the nearest surface oxygen atoms (s-O's), although we find that corresponding structures in which the cg-H atom(s) remain attached to the molecule(s), are found to be degenerate in energy (see Sec. S6). (b) A chain of dimers made of two intact molecules connected *via* a double H-bond at the respective carboxylic groups. This chain is by 1.23 eV less favorable (per dimer) than the one in (a) and its size goes well beyond two unit cells along [010].

These single molecule and dimer geometries provide a crucial preliminary information for computing the energy barriers characterizing single and double de-Ha processes, followed by the dimerization. In the following, we will use the NEB method to understand how the effective barrier reduction can occur by considering several possible scenarios. A comparison of the corresponding energy barriers will guide us in the identification of the most probable reaction mechanism.

Independent de-Ha's involving flat-lying molecules (Fig. 6): In the first scenario, we assume that two molecules dehalogenate independently and subsequently diffuse towards each other forming a dimer. Given the substantial equivalence of each de-Ha reaction, only one barrier involving a single IBA needs to be calculated. The most favorable process found among a set of reactions studied (see Sec. S7) is illustrated in Fig. 6. The IBA is chemisorbed, with the phenyl radical bound to the surface. The barrier associated is 1.76 eV, by 1.43 eV lower than the gas-phase value (3.19 eV). The reduction is due both to the chemisorption and to the adsorption energy of the iodine on the surface (clearly missing in the gas phase), which we calculated separately as 0.81 eV (see Sec. S4). This results in a much lower final state (FS) energy, which in turn lowers the barrier itself. Total energy comparisons indicate that the FS of a chemisorbed de-Ha IBA on calcite would be less stable than the intact molecule initial state (IS). This is in contrast with analogous processes on noble metal surfaces, where iodobenzene and bromobenzene dehalogenations were calculated to have a lower total energy of the products[26]. Dehalogenation processes of this kind happen independently for each molecule. Their energy cost is clearly reduced by the surface (with respect to the gas phase). These reactions must be followed by diffusion of de-Ha molecules and their dimerization, provided that they find themselves at a short distance. Iodine atoms might also diffuse and recombine forming I_2. Experimentally, we do not have evidence for iodine atoms or I_2 molecules adsorbed on the surface. However, we could easily envision that I_2 desorb and go to the gas-phase, a process that has an associated cost of 0.78

eV (see Sec. S4). As we are considering reactions at high temperatures, changes in the free energy rather than the total energy must be considered, including the gain in entropy. Here, the major contribution comes from the diatomic gas formed upon the I_2 desorption. We estimate this contribution in a range from 2.6 to 3.2 eV per I_2 molecule in a 500 to 600 K temperature interval (Section S11). This high contribution makes the re-combinative desorption of I_2 thermodynamically favourable. In summary, the sequence of independent mechanism(s) described above and involving chemisorption must be considered as a possible pathway, clearly pointing to a catalytic role of the surface in lowering the rate determining barrier associated to the de-halogenation step.

Importantly, the same I_2 gas entropy gain applies also to all FS considered below. The main fundamental process initiating the formation of the iodine molecule is the de-Ha reaction. Hence by comparing various de-Ha mechanisms we should be able to assess their likelihood. According to the classical transition state (TS) theory in the harmonic approximation[56], the transition rate depends on the energy barrier (the difference of total energies between the TS and the minimum), on temperature and on entropic effects. The latter basically enter via the pre-factor containing the ratio of vibrational frequencies in the minimum and at the saddle point. In most cases the pre-factor varies within 10^{12} and 10^{14} s^{-1}, corresponding to the characteristic vibrational frequency "responsible" for the transition. Numerically, however, the importance of the pre-factor is less severe than of the energy barrier, which enters exponentially. This enables us to concentrate on the calculation of energy barriers in assessing the likelihood of various transitions. In these calculations the pre-factor was fixed at the value of 10^{13} s^{-1}. Note that entropic effects due to the I_2 gas would only affect the free energy of the final state and hence will not affect the calculated transition rates.

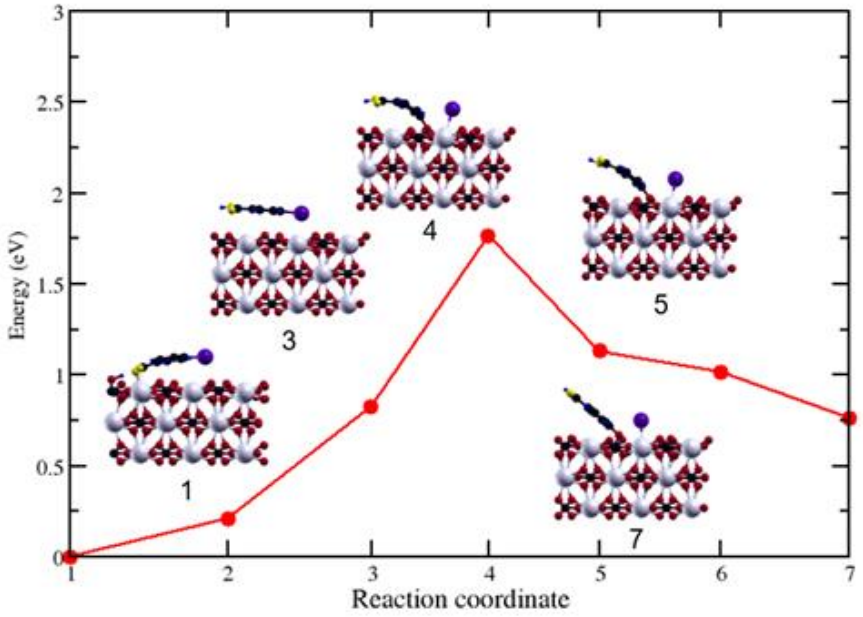


Figure 6. Minimum energy profile and reaction path for a single dehalogenation process involving one IBA molecule flat-lying on calcite (10.4). The transition between the IS (Fig. 3(b)) and FS (Fig. 3(e)) with the energy barrier of 1.76 eV. The barrier is reduced with respect to the gas-phase due to the stabilization effect achieved by the de-Ha IBA chemisorption and by the iodine adsorption in FS. Note that for a process having as IS an upright IBA (not shown), a barrier like the gas phase one is expected (3.19 eV), as the iodine would have to go into the gas phase first.

Two flat-lying molecule processes (Fig. 7): As a second possible mechanism, we consider two molecules which react together in a *unique* process. This involves a double dehalogenation with subsequent dimerization. Several cases are considered, all sharing energetically very similar IS and the same FS. As an IS, we consider two separate flat-lying molecules facing each other at some distance along [010]. Each unreacted molecule is in a configuration similar to the one of Fig. 3(b), the second most stable along [010]. The FS is the stable configuration of a covalent dimer also oriented along [010], with the iodine atoms forming a iodine molecule I_2 in the gas phase (Fig. 5 and Sec. S6). This dimer is the building block of the observed chains along [$\bar{4}\bar{2}1$]. The reaction is exothermic, with a total energy gain of $\Delta E_{(FS-IS)}$ = -1.14 eV.

The first process is illustrated in Fig. 7(a), where two molecules belonging to a chain dehalogenate without chemisorption on the surface. The barrier is 2.8 eV. Note that this value is much lower than the sum of two concomitant iodine detachments, as the NEB identifies an *asynchronous* pathway where the two iodine atoms detach at different times. This is, in general, another important aspect that facilitates the reaction. As we have seen above, also the IBA chemisorption after dehalogenation plays an important catalytic role. In Fig. 7(b) we consider a second process where the dehalogenations are asynchronous and both phenyl radicals bind to calcite, individually realizing the most stable dehalogenated configuration on the surface (similar to Figs. 3(d,e)). The calculated barrier is 2.8 eV, still rather high and similar to the one in Fig. 7(a). This shows that in the whole process involving *two* molecules, the IBA's chemisorption has a marginal catalytic effect, even if in the independent scenario involving single molecules helped to lower the barrier. Interestingly, what makes this process still costly are the synchronous detachments of the phenyl radicals from the surface before dimerization, as shown in Fig. 7(b).

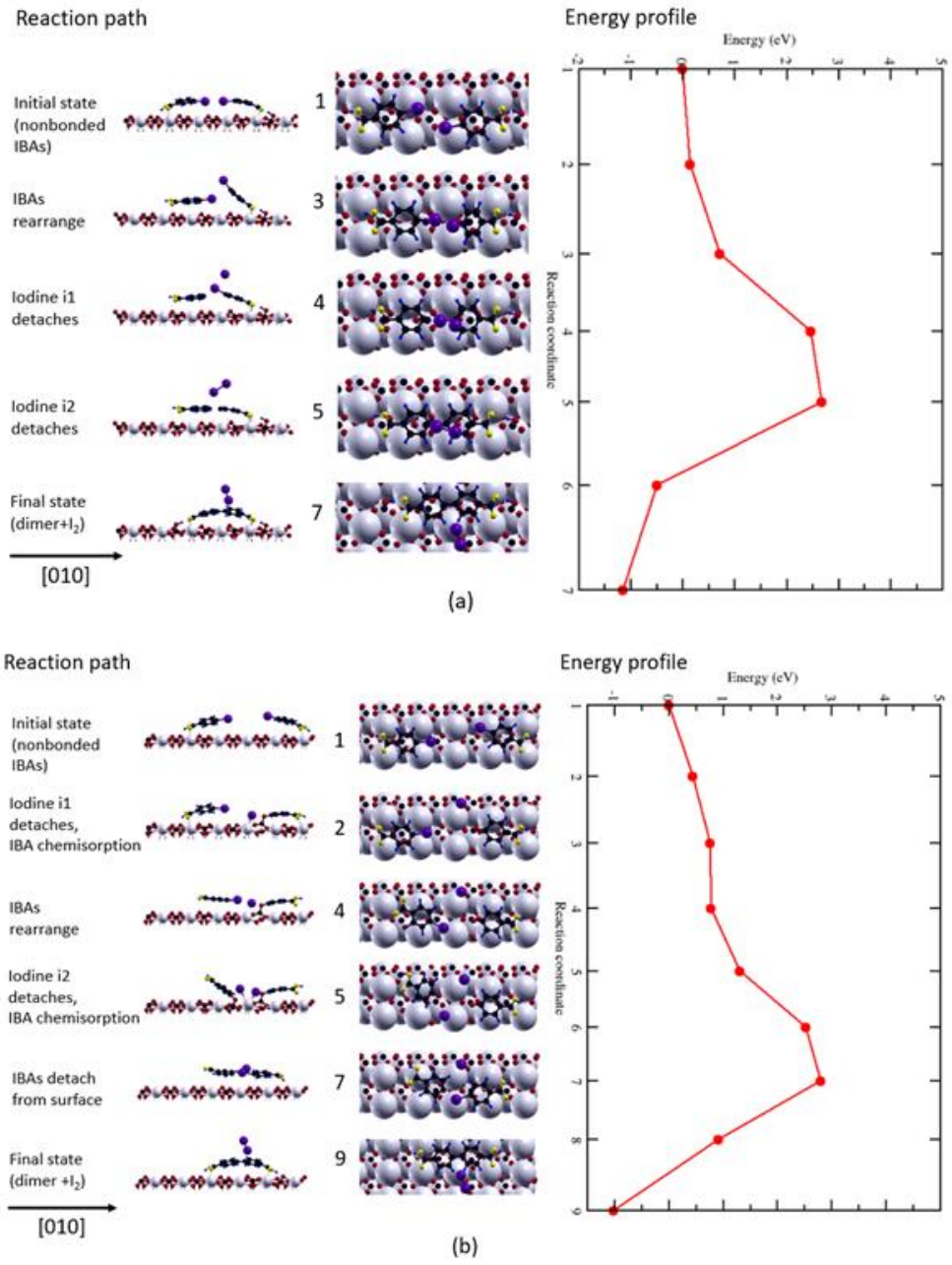


Figure 7. Reaction path and minimum energy profile of two different processes involving two flat-lying molecules oriented along the [010] direction on calcite [55]. In both cases the reaction consists of two dehalogenation reactions followed by a dimerization and formation of the I_2 molecule. All reactions are exothermic with the total energy gain of $\Delta E_{(FS-IS)} \sim -1$ eV. (a) A process where the dehalogenations are asynchronous, with the first "assisted" by the second, with a barrier of 2.8 eV. (b) In this process the de-Ha phenyl radicals bind to the surface in the

intermediate states (R_c=2-5) before dimerization. The barrier is 2.8 eV, similar to (a). The main peak at R_c=6,7 is related to the two synchronous detachments from the surface that are required for the de-Ha IBA's to dimerize.

Two molecules mechanism with drastically reduced energy barriers (Fig. 8). It has been mentioned above that the upright phase is the molecular arrangement observed just before the formation of the chains of dimers (Fig. 2). Therefore, as the third possible mechanism, we consider the IS consisting of two *upright* molecules anchored on calcite with the carboxylate group oriented along $[\bar{4}\bar{2}1]$ and placed apart along [010]. Their H atoms are initially detached to the respective nearest surface O atoms, as tentatively concluded from the experimental results [32]. This geometry facilitates the formation of a dimer in the relaxed chain geometry of Fig. 5(a). At the same time, we believe that this process must be extremely relevant for the phase transition under discussion, as this IS (even though the upright *single* molecule is less stable than the flat-lying one) can be a possible geometry within the upright phase, e.g., at the edge of an island, or during the structural changes happening at 580K. Hence, an IS made of upright molecules is likely to bring to the IBA dimer and, in our simplified scheme, the corresponding process would be one of the elementary steps of the phase transition between upright phase and chains of dimers.

In Fig. 8 we present a rather non-trivial reaction mechanism that incorporates most of the catalyzing factors highlighted above. During the reaction (Fig. 8, left panel), the cost of the dehalogenation involving the first iodine atom (*i1*) is reduced, as the second (*i2*) is first "exchanged" and then "shared" between the two molecules. Thus, a plateau of intermediate low energy states with both molecules attached to the same iodine (*i2*) is formed (Fig. 8, right panel). Only at that point *i2* detaches and joins *i1* to form the I_2 molecule, while the IBA's dimerize *via* the covalent linking between phenyls. The detachments give rise to two barriers whose height (1.85 eV and 1.5 eV) is drastically reduced in comparison with all two-molecule processes

considered above, whose barriers are ~2.8 eV. Note that also in this case the barrier for two de-Ha's is *far less* than the sum of individual barriers (if they happened independently in the gas phase) as each detachment is somehow "assisted" by the presence of the other iodine, in a cooperative mechanism to lower the barrier. Similar to the mechanisms presented in Fig. 7, the reaction is exothermic, with an energy difference $\Delta E_{(FS-IS)}$ = ~ -1 eV (the backward barrier is ~2.6 eV). Moreover, the reaction products would be stabilized further by the I_2 gas entropic contribution discussed above. This specific lower barrier mechanism is based on the following key factors facilitating the reaction under study: *i)* two asynchronous dehalogenations; *ii)* exchange and sharing of the iodine atoms during the detachments from the molecules; *iii)* I_2 recombination. We note that the mechanism in Fig. 8 implies a barrier reduction even if the molecules are considered in the gas phase, with their carboxylic groups kept in fixed positions. The corresponding calculated barrier is 1.9 eV (see Sec. S9), only slightly higher than for the process on the substrate. However, the presence of the calcite substrate is still crucial as a two-dimensional support providing the conditions for the reaction to happen. Namely the possibility for the molecules to stand upright and well anchored to the surface (see also the discussion of the *dense* phase in Ref. 43). These conditions, we emphasize, would be hardly realized in the three-dimensional gas phase, making the reaction very unlikely.

Finally, it was speculated[4,19] that the cleavage of the iodine-phenyl bond could be facilitated by the deprotonation of the IBA carboxylic acid group on calcite (re-interpreted as an H-bond formation[43]). This interesting idea is based on the consideration that the molecule, in a negative state of charge, would "lose" more easily the iodine atom. We made a detailed analysis to address this idea. We compared the dissociation energies from a "deprotonated" IBA with the one from a neutral molecule, in the gas phase. We also compared the de-Ha energies starting from configurations on the surface in Figs. 4(a) (where the cg-H is attached to the molecule) and Fig. 4(b) (where the H-bond is realized). Our calculations do not support this idea: we find that

in the gas phase the de-Ha energy from a neutral molecule is the smallest (3.19 eV, see Tab. S1), while on the surface the de-Ha energies from the structures in Fig. 4(a) and 4(b) are very similar (3.19 eV and 3.20 eV, respectively) and essentially equal to the gas phase value.

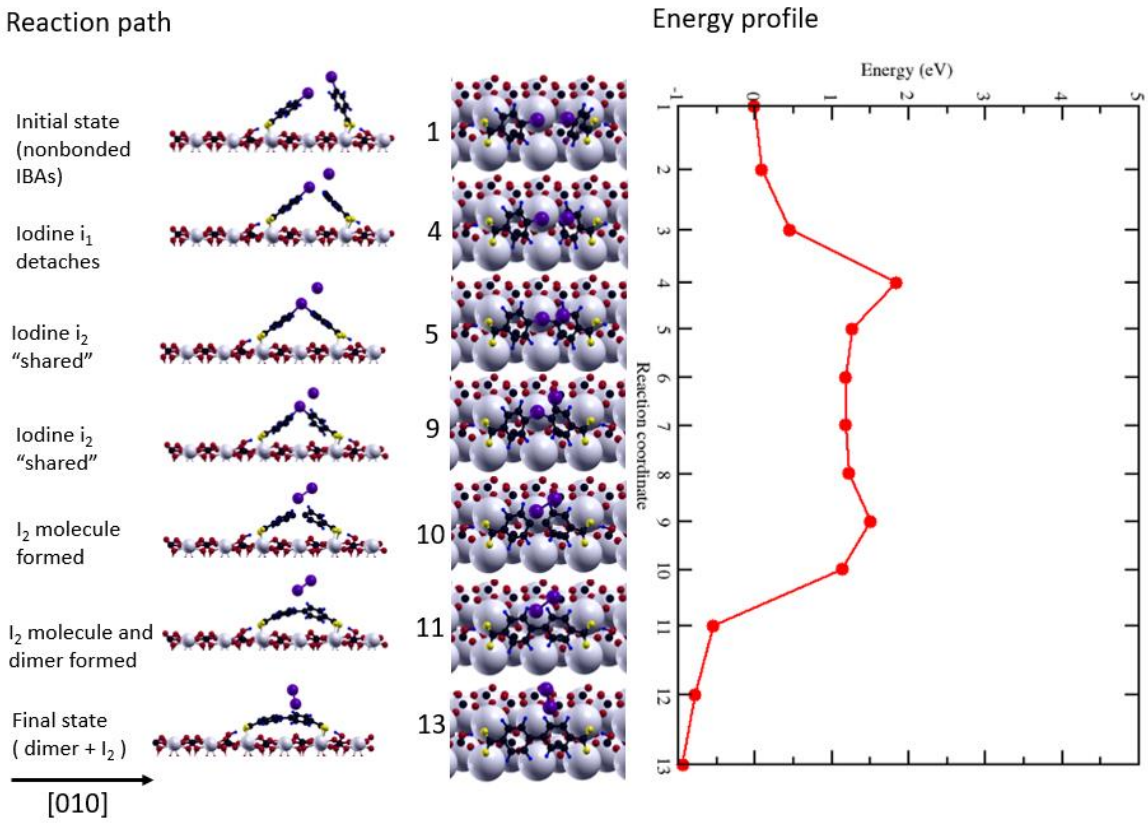


Figure 8. Reaction path and minimum energy profile of the most favorable process of double dehalogenation and dimerization of two IBA molecules on calcite (10.4). The initial state (R_c=1) shows two upright non-bonded molecules. The first barrier (R_c=4) is due to the dehalogenation involving the first iodine (*i1)*, followed by a sequence of states that "assist" this process (reduce its energy cost) and result in the second iodine (*i2)* to first exchange (R_c=5) and then become shared between the molecules (R_c=6-9). The second barrier (R_c= 9-10) stems from the second dehalogenation, "assisted" by the immediate I_2 formation (R_c=10) and the covalent dimerization (R_c=11-13). The energy gain relative to the initial state is ~1 eV, indicating the exothermic character of the reaction.

3. Conclusion

In this work, we employed *ab initio* theory and NC-AFM experiments to analyze the formation mechanisms of covalently linked molecular dimers that assemble in one-dimensional chains on an insulating calcite (10.4) substrate. The dimers are made of *4-iodobenzoic acid* (IBA) organic molecules, covalently connected via a (double) dehalogenation reaction. The latter is nominally very costly to be realized in the gas phase at the temperatures where the formation of this structure on calcite takes place. Focusing on this reaction and on the subsequent dimerization process, we use density functional theory and Nudged Elastic Band method to investigate in detail several relevant reaction paths and identify the key factors that allow these processes to happen on the substrate, with the resulting covalent assembly and chain formation.

The central result of our work is the identification of two very different processes where the dehalogenation barrier is strongly reduced with respect to the case of a single molecule in the gas phase (3.19 eV), due to the catalytic role of the surface. The first involves *single* molecules that dehalogenate independently and lower the barrier (1.76 eV) by chemisorbing on the surface upon losing the iodine atom. The second is a non-trivial and distinct reaction mechanism where an "assisted" double dehalogenation followed by I_2 formation and IBA dimerization is found as a minimum energy path. During the process, the two iodines dehalogenate in an asynchronous way and cooperate with each other to optimize the cost of their molecular detachment, prior to dimerization. Also this specific mechanism allows a drastic reduction of the energy barrier to 1.85 eV. We find that the surface does not play a direct role in this specific mechanism and in facilitating the reaction involving two IBA molecules. However, it provides the suitable sites for the molecule-substrate anchoring (via Ca-O bonds), which is necessary to form the upright phase from which the dimer chains are formed. The surface acts then as a two-dimensional support necessary to initiate the non-trivial process found.

To the best of our knowledge this is the first theoretical study of covalent dimerization mechanisms on insulators. The identification of such mechanisms helps to realize robust and stable molecular architectures electronically decoupled from the substrate, which makes them particularly suitable for charge transport applications. It is also an important step to increase the molecule/substrate combinations that allow a proper functionalization of bulk insulating surfaces with covalent networks, a very challenging task to date.

COMPUTATIONAL METHODS

Density functional theory (DFT) and Nudged Elastic Band (NEB) calculations were performed with the planewave-pseudopotential package Quantum ESPRESSO [47], using Ultrasoft pseudopotentials [48] with a wave function (charge) energy cutoff of 408 eV (4080 eV) and a GGA-PBE [49] exchange-correlation functional. The Grimme-D2 van der Waals interaction [50] was included. All calculations presented, including the ones with the substrate, were performed using spin-resolved DFT. Brillouin-zone sampling included the $\mathbf{k} = \Gamma$ point only. The calcite (10.4) substrate was modeled with a periodically repeated slab of three layers, allowing a vacuum gap between the adsorbed molecular species and the bottom layer of the slab replica of ~20 Å. Molecule atoms and surface atoms belonging to the first two layers were allowed to relax. Forces were relaxed up to 0.026 eV/Å, with a 1.36×10^{-7} eV cut-off on the total energy. A smearing of 1.36×10^{-2} eV was used to improve convergence in the electronic iterations. Energy and force parameters have been chosen to make sure that the calculations are well converged. Convergence of NEB processes with respect to the number of images was checked in a few cases, and the used number of images was sufficient in all checked cases. The climbing image method [51] was used in all NEB calculations.

EXPERIMENTAL METHODS

The experiments were performed under ultra-high vacuum (UHV) conditions with a base pressure typically better than 10^{-10} mbar. Optical quality calcite ($CaCO_3$) crystals (Korth Kristalle GmbH, Kiel, Germany) were cleaved *in situ* prior to the deposition of 4-iodobenzoic acid (IBA, Aldrich, Munich, Germany). Subsequently, molecularly resolved images were taken with an Omicron Nanotechnology (Taunusstein, Germany) variable-temperature atomic force microscope (AFM), operated in the frequency modulation non-contact (NC) mode. Further details can be found in [4].

Acknowledgments

C. G. and Y. W. gratefully acknowledge financial support from the National Natural Science Foundation of China under Grant No. 11574238. AF and LK thank the EPSRC grant (EP/J019844/1) for funding. Via our membership of the UK's HEC Materials Chemistry Consortium, which is funded by EPSRC (EP/L000202), this work used the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk). AK gratefully acknowledges financial support from the EU through grant PAMS (seventh framework program GA 610446). C.G., AF and LK would like to thank Chiara Paris for help in providing some useful material for the realization of this work.

Supporting Information available: 1) IBA molecule in the gas phase: structure and dehalogenation processes. 2) IBA adsorption geometries on calcite (10.4). 3) Dehalogenated IBA molecule with a iodine atom on calcite (10.4). 4) Iodine atom and I_2 molecule adsorption geometries on calcite (10.4). 5) H-bond energy. 6) IBA dimer adsorption geometries on calcite (10.4). 7) Single molecule dehalogenation on calcite (10.4). 8) Dehalogenation process with I_2

molecule formation. 9) Double dehalogenation-dimerization process in the gas phase. 10) Two non-bonded IBA molecules adsorption geometries on calcite (10.4). 11) Entropy contribution of a gas of I_2 molecules to the free energy. 12) Spin-resolved versus spin-unresolved energy barriers in processes including the substrate. 13) Stable configurations of two hydrogen atoms in the presence of a calcite (10.4) surface.

References

[1] Matena, M.; Riehm, T.; Stöhr, M.; Jung, T. A.; Gade, L. H. Transforming surface coordination polymers into covalent surface polymers: Linked polycondensed aromatics through oligomerization of N-heterocyclic carbene Intermediates. *Angew. Chem., Int. Ed.* **2008**, *47,* 2414– 2417.

[2] Grill, L.; Dyer, M.; Lafferentz, L.; Persson, M.; Peters, M. V.; Hecht, S. Nano-architectures by covalent assembly of molecular building blocks. *Nat. Nanotechnol.* **2007**, *2*, 687– 691.

[3] Rahe, P.; Kittelmann, M.; Neff, J.L.; Nimmrich, M.; Reichling, M.; Maass, P.; Kühnle, A. Tuning molecular self-assembly on bulk insulator surfaces by anchoring of the organic building blocks. *Adv. Mater.* **2013**, *25,* 3948–3956.

[4] Kittelmann, M; Rahe, P.; Nimmrich, M; Hauke, C. M.; Gourdon, A; Kühnle, A. On-surface covalent linking of organic building blocks on a bulk insulator. *ACS Nano* **2011**, *5*, 8420–8425.

[5] Floris, A.; Haq S.; In't Veld, M.; Amabilino, D. B.; Raval, R; Kantorovich, L. Driving forces for covalent assembly of porphyrins by selective C-H bond activation and intermolecular coupling on a copper surface, *J. Am. Chem. Soc.* **2016**, *138*, 5837–5847.

[6] Lindner, R; Kühnle, A. On-Surface Reactions, *ChemPhysChem* **2015,** *16*, 1582–1592.

[7] Björk, J. Reaction mechanisms for on-surface synthesis of covalent nanostructures *Journal of Physics: Condensed Matter* **2016**, *28,* 083002.

[8] In't Veld, M.; Iavicoli, P.; Haq S.; Amabilino, D. B.; Raval, R. Unique intermolecular reaction of simple porphyrins at a metal surface gives covalent nanostructures. *Chem. Commun.* **2008**, 1536-1538.

[9] Abdel-Mottaleb, M.M.S.; Gomar-Nadal, E.; Surin, M.; Uji-I, H.; Mamdouh, W.; Veciana, J.; Lemaur, V.; Rovira, C.; Cornil, J.; Lazzaroni, R.; Amabilino, D.B.; De Feyter, S.; De Schryver, F.C. Self-assembly of tetrathiafulvalene derivatives at a liquid/solid interface—compositional and constitutional influence on supramolecular ordering. *J. Mater. Chem*. **2005**, *15*, 4601-4615.

[10] Puigmartí-Luis, J.; Minoia, A.; Uji-I, H; Rovira, C; Cornil, J; De Feyter, S; Lazzaroni, S; Amabilino, D.B. Noncovalent control for bottom-up assembly of functional supramolecular wires. *J. Am. Chem. Soc.* **2006**, *128*, 12602-12603.

[11] Abdurakhmanova, N.; Floris, A.; Tseng, T.C.; Comisso, A.; Stepanow, S.; De Vita, A.; Kern, K. Stereoselectivity and electrostatics in charge-transfer Mn- and Cs-$TCNQ_4$ networks on Ag(100), *Nat. Comm.* **2012**, *3*, 940.

[12] Floris, A.; Comisso A.; De Vita, A. Fine-tuning the electrostatic properties of an alkali-linked organic adlayer on a metal substrate. *ACS Nano* **2013**, *7*, 8059–8065.

[13] Blüm M.-C.; Ćavar E.; Pivetta M; Patthey F.; Schneider W.-D. Conservation of chirality in a hierarchical supramolecular self-assembled structure with pentagonal symmetry. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 5334-5337.

[14] Tomba, G.; Stengel, M.; Schneider, W.-D.; Baldereschi, A.; De Vita, A. Supramolecular self-assembly driven by electrostatic repulsion: The 1D aggregation of rubrene pentagons on Au(111). *ACS Nano* **2010**, *4*, 7545-7551.

[15] Della Pia, A.; Riello, M.; Floris, A.; Stassen, D.; Bonifazi, D.; Tim S. Johnes.; A. De Vita.; Costantini, G. Anomalous coarsening driven by reversible charge transfer at metal organic interfaces. *ACS Nano* **2014**, 8, 12356–12364.

[16] Willner, A.E.; Byer, R.L.; Chang-Hasnain, C.J.; Forrest, S.R.; Kressel, H.; Kogelnik, H.; Tearney, G.J.; Townes, C.H.; Zervas, M.N., "Optics and Photonics: Key Enabling Technologies," in *Proceedings of the IEEE*, vol.100, no. Special Centennial Issue, **2012**, 1604-1643.

[17] Nilius, N.; Risse, T.; Schauermann, S.; Shaikhutdinov, S.; Sterrer, M.; Freund, H.-J. Model studies in catalysis. *Top. Catal.* **2011**, *54*, 4–12.

[18] Abel, M; Clair, S; Ourdjini, O; Mossoyan, M; Porte, L. Single layer of polymeric fe-phthalocyanine: An organometallic sheet on metal and thin insulating film. *J. Am. Chem. Soc.* **2011**, *133*, 1203–1205.

[19] Kittelmann, M.; Nimmrich, M.; Lindner, R.; Gourdon, A.; Kühnle, A. Sequential and site-specific on-surface synthesis on a bulk insulator. *ACS Nano* **2013**, *7*, 5614-5620.

[20] Lindner, R,; Rahe, P.; Kittelmann, M.; Gourdon, A.; Bechstein, R.; Kühnle, A. Substrate Templating Guides Photo-Induced Reaction of C60 on Calcite. *Angew. Chem. Int. Ed.* **2014**, *53*, 7952-7955.

[21] Haq, S.; Hanke, F.; Sharp, J.; Persson, M.; D. B. Amabilino.; Raval, R. Versatile bottom-up construction of diverse macromolecules on a surface observed by scanning tunneling microscopy. *ACS Nano* **2014**, *8*, 8856-8870.

[22] Lafferentz, L.; Eberhardt, V.; Dri, C.; Africh, C.; Comelli, G.; Esch, F.; Hecht, S.; Grill, L. Controlling on-surface polymerization by hierarchical and substrate-directed growth. *Nature Chem.* **2012,** *4*, 215-220.

[23] Cai, J. M.; Ruffieux, P.; Jaafar, R.; Bieri, M.; Braun, T.; Blankenburg, S.; Muoth, M.; Seitsonen, A. P.; Saleh, M.; Feng, X. L. et al., Atomically Precise Bottom-up Fabrication of Graphene Nanoribbons, *Nature* **2010**, *466*, 470– 473.

[24] Lee, J.; Dougherty DB.; Yates J.T. Jr. Chemisorbed benzoate-to-benzene conversion via phenyl radicals on Cu(110): kinetic observation of conformational effects. *J. Am. Chem. Soc.* **2006**, *128*, 6008-6009.

[25] Treier, M.; C A Pignedoli.; Laino, T.; Rieger, R.; Müllen, K.; Passerone, D. ; Fasel, R.. Surface-assisted cyclodehydrogenation provides a synthetic route towards easily processable and chemically tailored nanographenes. *Nature Chem.* **2010**, *3*, 61–67.

[26] Björk, J.; Hanke, F.; Stafström, S., Mechanisms of halogen-based covalent self-assembly on metal surfaces. *J. Am. Chem. Soc.* **2013**, *135*, 5768–5775.

[27] Haq, S.; Hanke, F.; M. S Dyer; Persson, M.; Iavicoli, P.; D. B. Amabilino; Raval, R. Clean coupling of unfunctionalized porphyrins at surfaces to give highly oriented organometallic oligomers. *J. Am. Chem. Soc.* **2011**, *133*, 12031-12039.

[28] Hanke, F.; Haq, S.; Raval, R.; Persson, M. Heat-to-connect: Surface commensurability directs organometallic one-dimensional self-assembly. *ACS Nano* **2011**, *5*, 9093-9103.

[29] Björk, J.; Zhang, Y. Q.; Klappenberger, F.; Barth J. V.; Stafström, S. Unraveling the mechanism of the covalent coupling between terminal alkynes on a noble metal. *J. Phys. Chem.* C **2014**, *118***,** 3181-3187.

[30] Gao H. Y.; Franke, J.; Wagner, H.; Zhong, D.; Held P A.; Studer, A.; Fuchs, H. Effect of metal surfaces in on-surface glaser coupling. *J. Phys. Chem.* C **2013**, *117*, 18595-18602.

[31] Björk, J.; Stafström, S.; Hanke, F. Zipping up: Cooperativity drives the synthesis of graphene nanoribbons. *J. Am. Chem. Soc.* **2011**, *133*, 14884-14887.

[32] Kittelmann, M.; Nimmrich, M.; Neff, J. L.; Rahe, P.; Greń, W.; Bouju, X.; Gourdon, A.; Kühnle, A. Controlled activation of substrate templating in molecular self-assembly by deprotonation. *J. Phys. Chem. C* **2013**, *117*, 23868-23874.

[33] Pawlak, R.; Nony, L.; Bocquet, F.; Olson, V.; Sassi, M.; Debierre, J.M.; Loppacher, C.; Porte, L. Supramolecular assemblies of 1,4-benzene diboronic acid on KCl(001). *J. Phys. Chem. C* **2010**, *114*, 9290-9295.

[34] Rahe, P.; Lindner, R.; Kittelmann, M.; Nimmrich, M., Kühnle, A. From dewetting to wetting molecular layers: C_{60} on $CaCO_3$(10.4) as a case study. *Phys. Chem. Chem. Phys.* **2012**, *14*, 6544-6548.

[35] Kunstmann, T.; Schlarb, A.; Fendrich, M.; Wagner, T.; Möller, R.; Hoffmann, R. Dynamic force microscopy study of 3,4,9,10-perylenetetracarboxylic dianhydride on KBr(001). *Phys. Rev. B* **2005**, *71*, 121403.

[36] Nony, L.; Bennewitz, R.; Pfeiffer, O.; Gnecco, E.; Baratoff, A.; Meyer, E.; Eguchi, T.; Gourdon, A., C. Joachim. Cu-TBPP and PTCDA molecules on insulating surfaces studied by ultra-high-vacuum non-contact AFM. *Nanotechnology* **2004**, *15*, S91.

[37] Neff, J. L.; Kittelmann, M.; Bechstein, R.; Kühnle, A. Decisive influence of substitution positions in molecular self-assembly. *Phys. Chem. Chem. Phys.* **2014**, *16*, 15437-15443.

[38] Kittelmann, M.; Rahe, P.; Gourdon, A.; Kühnle, A. Direct visualization of molecule deprotonation on an insulating surface. *ACS Nano* **2012**, *6*, 7406-7411.

[39] Kittelmann, M.; Rahe, P.; Kühnle, A. Molecular self-assembly on an insulating surface: Interplay between substrate templating and intermolecular interactions. *J. Phys.: Condens. Matter* **2012**, *24*, 354007.

[40] Hauke, C. M.; Bechstein, R.; Kittelmann, M.; Storz, C.; Kilbinger, A. F.; Rahe, P.; Kühnle, A. Controlling Molecular Self-Assembly on an Insulating Surface by Rationally Designing an Efficient Anchor Functionality That Maintains Structural Flexibility, *ACS Nano* **2013**, *7*, 5491.

[41] Rahe, P.; Nimmrich, M.; Greuling, A.; Schütte, J.; Stará, I. G.; Rybáček, J.; Huerta-Angeles, G.; Starý, I.; Rohlfing, M.; Kühnle, A. Toward molecular nanowires self-assembled on an insulating substrate: Heptahelicene-2-carboxylic acid on calcite (10.4). *J. Phys. Chem. C* **2010**, *114*, 1547-1552.

[42] Neff, J.L.; Söngen, H.; Bechstein, R.; Maass, P.; Künle, A. Long range order induced by intrinsic repulsion on an insulating substrate, *J. Phys. Chem. C* **2015**, *119*, 24927.

[43] Paris, C.; Floris, A.; Aeschlimann, S.; Kittelmann, M.; Kling, F.; Bechstein, R.; Kühnle, A. Increasing the templating effect on a bulk insulator surface: From a kinetically trapped to a thermodynamically more stable structure, *J. Phys. Chem. C* **2016**, *120 (31)*, 17546-17554.

[44] Jónsson, H.; Mills, G.; Jacobsen, K. W. Nudged elastic band method for finding minimum energy paths of transitions, in classical and quantum dynamics in condensed phase simulations, Ed. B. J. Berne, G. Ciccotti and D. F. Coker, 385 (World Scientific, **1998**).

[45] Lide, D. R., Bond strengths in polyatomic molecules. In CRC Handbook of Chemistry and Physics; (ed.) CRC Press: Boca Raton, FL, **2005.**

[46] Morchutt, C.; Björk, J.; Krotzky, S.; Gutzler, R.; Kern, K. Covalent coupling via dehalogenation on Ni(111) supported boron nitride and graphene. *Chem. Commun.* **2015**, *51***,** 2440-2443.

[47] Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I. et al. QUANTUM ESPRESSO: a modular and open-source software project fo quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**,

[48] Vanderbilt, D., Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **1990**, *41*, 7892.

[49] Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett*. **1996**, *77*, 3865–3868.

[50] Grimme, S. Semiempirical gga-type density functional constructed with a long-range dispersion correction. *J. Comp. Chem.* **2006**, *27*, 1787-1779.

[51] Henkelman, G.; Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113*, 9901-9904.

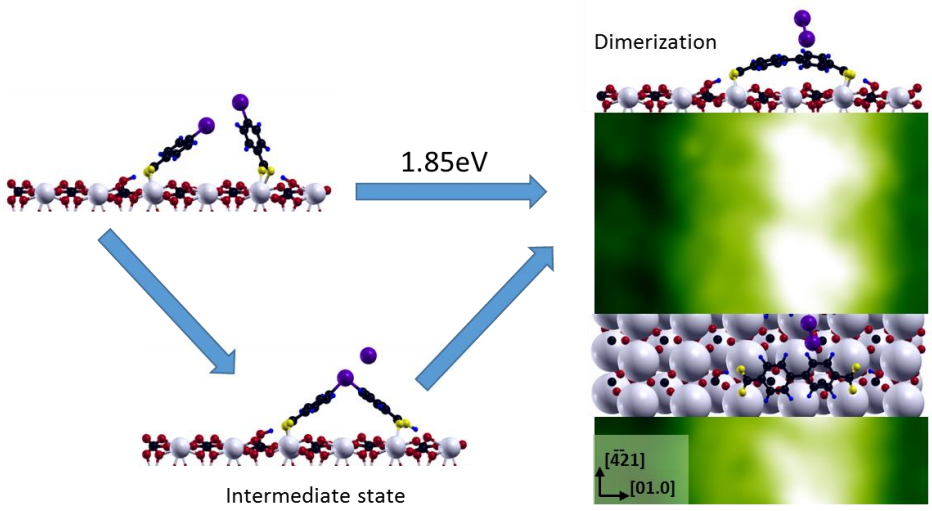
[52] For more clarity we denote by "chemisorption" only the process where a bond between the phenyl radical and an oxygen of the surface is realized after the dehalogenation.

[53] A rough estimate of the temperature necessary for this reaction to happen can be obtained by assuming a simple Arrhenius law. Taking an attempt frequency of 10^12 s^{-1}, the energy barrier of ~3 eV corresponds to a temperature of nearly 1000 K when assuming that half of the molecules shall be dissociated after 30 min. This is surely way more than the experimentally required temperature.

[54] This is corroborated by our NEB calculations showing, in some cases, that the H atom shuttles during the process, indicating that a very low energy barrier between the two states must exist. Note also that the effects of the H position on the de-halogenation and dimerization barriers is rather marginal.

[55] We do not define the reaction coordinate here as it is given by a linear combination of many atomic degrees of freedom to indicate the direction to the transition state. Such definition can only be meaningful within a narrow vicinity of each point along the minimum energy path. This approach is sufficient for our purposes since we are exclusively interested in the barriers associated with different paths. In this case it is also possible showing several paths on one graph even so they do not correspond to the same reaction coordinates.

[56] Vineyard, G.H. Frequency factors and isotope effects in solid state rate processes. *J. Phys. Chem. Solids* **1957** *3*, 121-127.



TOC Graphic